Computers & Geosciences 52 (2013) 246–257

Contents lists available at SciVerse ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Neural network modeling and prediction of resistivity structures using VES Schlumberger data over a geothermal area

Upendra K. Singh [a,*], R.K. Tiwari [b], S.B. Singh [b]

[a] Department of Applied Geophysics, Indian School of Mines, Dhanbad-826 004, India
[b] National Geophysical Research Institute, Uppal Road, Hyderabad-500 006, India

ARTICLE INFO

Article history:
Received 30 January 2012
Received in revised form
11 September 2012
Accepted 18 September 2012
Available online 22 October 2012

Keywords:
ANN
Learning parameters
VES
Inversion
Geothermal area

ABSTRACT

This paper presents the effects of several parameters on the artificial neural networks (ANN) inversion of vertical electrical sounding (VES) data. Sensitivity of ANN parameters was examined on the performance of adaptive backpropagation (ABP) and Levenberg–Marquardt algorithms (LMA) to test the robustness to noisy synthetic as well as field geophysical data and resolving capability of these methods for predicting the subsurface resistivity layers. We trained, tested and validated ANN using the synthetic VES data as input to the networks and layer parameters of the models as network output. ANN learning parameters are varied and corresponding observations are recorded. The sensitivity analysis of synthetic data and real model demonstrate that ANN algorithms applied in VES data inversion should be considered well not only in terms of accuracy but also in terms of high computational efforts. Also the analysis suggests that ANN model with its various controlling parameters are largely data dependent and hence no unique architecture can be designed for VES data analysis. ANN based methods are also applied to the actual VES field data obtained from the tectonically vital geothermal areas of Jammu and Kashmir, India. Analysis suggests that both the ABP and LMA are suitable methods for 1-D VES modeling. But the LMA method provides greater degree of robustness than the ABP in case of 2-D VES modeling. Comparison of the inversion results with known lithology correlates well and also reveals the additional significant feature of reconsolidated breccia of about 7.0 m thickness beneath the overburden in some cases like at sounding point RDC-5. We may therefore conclude that ANN based methods are significantly faster and efficient for detection of complex layered resistivity structures with a relatively greater degree of precision and resolution.

1. Introduction

Vertical electrical sounding (VES) is one of the most popular non-seismic methods generally used for the resource (i.e. ground water, mineral, geothermal etc.) exploration and geo-hydrological investigations. Physics-based numerical model has been commonly used for characterization of subsurface geological structures. The data are usually arranged and contoured in the form of pseudo-section, which gives an approximate picture of the subsurface resistivity structure. However, this objective, in its entirety, is somewhat difficult to achieve in the areas with the complex geology settings (Griffith and Barker, 1993), convoluted signals and non-linear relationship between the measured data. One of the conventional techniques that have been applied very frequently to overcome this problem is trial and error method. Trial and error method uses the concept of finite difference and finite element schemes and has been successful to provide considerably better results. However, these approaches are also time consuming, and computationally expensive. Furthermore, predictable discrepancies between the predicted model and true model exist, which reduce the simulation and prediction accuracy. The model obtained could also suffer from interpreter bias (Sasaki, 1989). Because of these recognized limitations, it is imperative to search for an alternative modeling perspective of geo-electrical data interpretation, which allows one to infer the spatial distribution of subsurface resistivity from the observed apparent resistivity data (resistivity inverse problem). Recently, artificial neural network (ANN) techniques have emerged as one of the efficient tools for geophysical data processing, e.g., prediction, inversion, feature classification, data compression, etc. (Calderon-Macias et al., 2000; El-Qady and Ushijima, 2001; Poulton et al., 1992; Raiche, 1991; Roth and Tarantola, 1994; Singh et al., 2005, 2006, 2010; Van der Baan and Jutten, 2000; Zhang et al., 2002; Zhang and Zhou, 2002). This is due to the fact that ANNs based techniques do not essentially require explicit characterization and accurate representation of the governing

* Corresponding author. Tel.: +91 326 2235606; fax: +91 326 2296563.
E-mail address: upendra_bhu1@rediffmail.com (U.K. Singh).

0098-3004/$ - see front matter
http://dx.doi.org/10.1016/j.cageo.2012.09.018

physical laws (Poulton et al., 1992). This method learns the behavior of the data pattern through its proper architecture design and training, and thereby renders better understanding. Hence, because of these non-rigid physical and mathematical framework and superior prediction performance, the ANNs based techniques have been successfully applied in various field data (Rumelhart et al., 1986, Lippman, 1987; Roth and Tarantola, 1994). Usage of ANNs has been proven useful in processing of 2D seismic data in hydrocarbon exploration (McCormack et al., 1993; Brown and Poulton, 1996).

In this paper, the main objective is to examine and understand the sensitivity of various processing elements i.e. controlling parameters (learning rate, momentum, number of hidden layers and number of neurons in hidden layers, weight and bias) on the ANNs inversion using the synthetic VES data. The selection of appropriate learning parameters, training and testing is vital for good performance of a network for processing actual measured data. In general, a large network and complicated input patterns require more training examples for an optimum generalization (Rajavelu et al., 1989). The optimum training set size, required for good generalization, was studied. Two kinds of networks were designed and implemented for training and testing of both 1-D and 2-D synthetic VES data sets. The effects of the network parameters, such as learning rate, momentum, input noise, number of hidden layers and, neurons in the hidden layer were monitored to understand the behavior of the network performance. The design of the network with aforementioned appropriate architectural parameters will be helpful in enhancing the clarity and robustness of actual VES data modeling. Finally ANN derived results were compared with obtained from conventional inversion method.

2. Brief theoretical background of artificial neural network

ANN is massively parallel-distributed processor that stores experimental knowledge and makes it available for use. It resembles the human brain in two respects: (i) knowledge is acquired through learning process by a network, and (ii) the inter-neuron connection strengths known as synaptic weights are used to store the knowledge (Haykin, 1994). A simple neural network as shown in Fig. 1 consists of three layers. The first (input) layer consists of neurons, each neuron of which receives one of the input variables. The intermediate (hidden) layer consists of neurons each of which computes a non-linear transformation as described below in Eq. (1). The third (output) layer consists of neurons, each of which computes a desired output variable. The output value of neurons depends on the weighted sum of an input and a weight. The sum may not fall on an optimal point of the curve, so one way to ensure

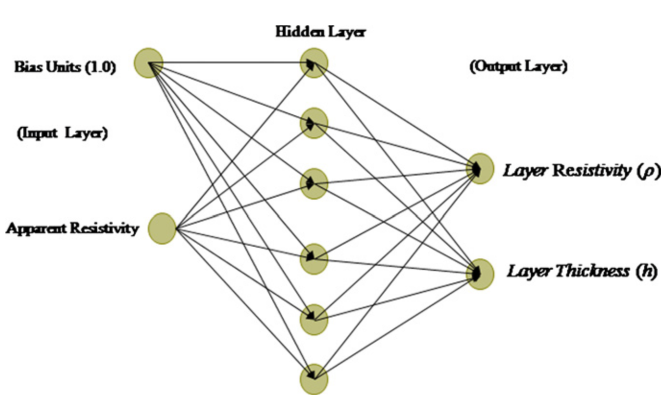

Fig. 1. Schematic diagram of supervised artificial neural network architecture.

that is to add a bias to the sum to shift it left or right along the curve.

This can be represented in mathematical term as follows:

$$y_j(N) = f(U_j(N)) = f\left(\sum_{i=0}^{n_i} w_{ji} x_i(N)\right) \quad \text{for} \quad j = 1, 2, 3, \ldots\ldots, n_j \tag{1}$$

$$z_k(N) = f(U_k(N)) = f\left(\sum_{j=0}^{n_j} w_{kj} y_j(N)\right) \quad \text{for} \quad k = 1, 2, 3, \ldots\ldots, n_k \tag{2}$$

where $x_i(N)$=input to neurons i of the input layer, $y_j(N)$=quantity computed by neuron j of the hidden layer and $z_k(N)$=output computed by neuron k of the output layer.

Parameter w_{ji} controls the strength of the connection between input neuron i and hidden neuron$j$, w_{kj} controls the strength of the connection between hidden neuron j and output neuron k. The commonly employed non-linear activation function was used.

$$\text{Sigmoid function}: \quad f(U_j) = \frac{1}{1+e^{-U_j}} \tag{3}$$

$$\text{Hyperbolic tangent function}: \quad f(U_j) = \frac{1-e^{-2U_j}}{1+e^{-2U_j}} \tag{4}$$

where $U_j \in [-\infty\ \infty]$ and $f(U_j)$ is bounded on (0, 1) for the sigmoid function and $f(U_j)$ is bounded on $(-1\ 1)$ for hyperbolic tangent (tan$h$) function. It is a continuous and bounded nonlinear transfer function; the sigmoid (logistic) and tan$h$ functions are most commonly used. A bias term is a processing element (PE) with a constant output value of 1.0 (Fig. 1). It is a trainable connection weight attached to each neuron in hidden and output layers (Widrow and Hoff, 1960). This helps to provide numerical stability.

ANN consists of numerous simple processing elements with large number of interconnections among them and is able to collectively solve complicated, especially dynamic non-linear problems. There are various kinds of ANN learning algorithms which are discussed in more detail elsewhere (Poulton, 2001) In this paper, we used supervised feed-forward ANN mode in a combination with backpropagation (BP) algorithm, adaptive backpropagation (ABP) and Levenberg-Marquardt based backpropagation algorithms (LMA) for solving VES inverse problems. In supervised feed-forward mode input/output pairs must be provided. Layers are connected to other layers in one direction from the input to the output layers. The output pairs are used for the comparison to the network computed output and differences are recorded for any changing parameters in the network. A brief description of the basic features is given, which will enable us to understand their applications more conveniently.

2.1. Backpropagation algorithm

The most widely used network is supervised neural network, which is based on the concept of BP algorithm (Haykin, 1994; Patterson, 1998). The BP is a first gradient descent method, which iterates the network learning process to adjust the weights in such way that the error becomes minimum between the network computed output and the observed data. The error is back propagated from output→hidden→input layer. In BP, the gradient vector of the error surface is calculated and the networks are updated by following equation (Govindaraju and Rao, 2000; Haykin, 1994)

$$\Delta w(N+1) = -\eta \tfrac{\partial f(w)}{\partial w} \big| w = w(N),$$
$$w(N+1) = w(N) - \eta \frac{\partial f(w)}{\partial w} \big| w = w(N), \tag{5}$$